Optimal Design of Pension Funds: A Mission Impossible?

JEROEN SUIJS[1] ANJA DE WAEGENAERE[2] PETER BORM[2]

March 20, 1998

Abstract

Nowadays many employers offer their employees the possibility of an insurance against too large losses in income when retiring or becoming disabled. This paper models the optimization problem of the employer when setting up such a so-called pension fund. not surprisingly, it turns out that the optimal solution depends on the premium the employees are willing to pay at most for the insurance. Since this is private information for an employee and hence not known to the employer, he needs to collect information regarding these maximum premiums. It is shown that in most cases the employer is unable to perfectly inform himself on these maximum premiums. So, he cannot create the right incentives for his employees to reveal their maximum premiums truthfully.

KEYWORDS: pension funds, insurance, Bayesian implementation.
JEL CODES: C72, D81, D82.

[1]CentER for Economic Research, Tilburg University, PO Box 90153, 5000 LE, Tilburg, The Netherlands.
[2]Department of Econometrics and CentER, Tilburg University, PO Box 90153, 5000 LE Tilburg, The Netherlands.

1 Introduction

Because of the growing population of elder people, social systems in many western European countries experience difficulties in guaranteeing an adequate income after retirement or disability. Especially those people with a higher income would experience a considerable loss of income when retiring or becoming disabled. Insurance companies anticipated this situation by developing several new products which insure people of an income of, say, 90% of their present income in case of disability, or an additional income above the pension supplied by the social system after retirement.

Employers on the other hand have to pay social premiums on the salaries paid to their employees. In the past decade, these premiums have gradually increased, and consequently, employers continuously look for other means to compensate their employees so as to pay less social premiums. A solution popular to many employers is creating a so-called 'pension fund' which, as the insurance companies do, insures employees against an excessive loss of income. The difference, however, is the fact that this fund is not an insurance company, but completely belongs to the employer's firm, who benefits from this construction due to a decreasing amount of social premiums that have to be paid.

In this paper we analyze the employer's decision problem with respect to the format of such a pension fund. To cover the future claims of the participating employees the employer establishes a pension fund. In our model we assume that any differences between the total sum of realized claims and the pension fund's total capital do not contribute to the employer's profit/loss. So, if the total of realized claims turns out to be lower than the pension fund's capital, then the profits of the employer do not increase. Similarly, if the total of realized claims turn out to be higher then the profits of the employer do not decrease. This means that the employer only benefits from the pension fund through a decrease in the social premiums that have to be paid. The pension fund, however, must be such that the probability that the total claims exceed the available capital is sufficiently low. So, roughly speaking, the employer's objective is to maximize the reduction in social premiums that have to be paid, subject to the earlier mentioned solvability constraint.

This paper focuses attention on the determination of the optimal insurance premium. This premium is the same for all insured employees. So, the employer is not allowed to discriminate between his employees. Reason for this assumption is that in some countrie, The Netherlands for example, price discrimination in this regard is illegal. For determining this insurance premium the employer lacks some vital information, namely, how much the employees want to pay at most for this insurance. In order to get this information the employer could simply ask each of his employees to tell him what insurance premium he wants to pay at maximum. Then, given this information, he could determine the optimal insurance premium. Acting in this way the employer implicitly assumes that his employees are honest and truthfully reveal their maximum premiums. The possibility of any strategic behavior by the employees is undeservingly ignored. The employees, however, could induce a lower insurance premium by not reporting the true maximum premium but another, lower one. Consequently, the employer would charge a lower premium than he would have charged in the optimum, resulting in less than optimal profits.

Summarizing, it is not as much the determination of the optimal insurance premium that poses any problems to the employer, but the acquisition of the correct information. To answer the question whether it is indeed possible for the employer to get the correct information or not we turn to implementation theory.

For surveys on implementation theory the reader is referred to *Moore (1992)* and *Palfrey (1992)*. Now, let us try to explain what implementation theory is all about. Consider the employer and his employees and suppose that $\overline{\pi}_i$ is the true maximum premium of employee i. If the employer exactly knows these premiums, he can determine the optimal premium π^*. Unfortunately he will not exactly know the premiums $\overline{\pi}_i$ in advance, so he cannot determine this optimal solution π^*. Instead, he can design a mechanism (or game) in which the employees participate, that yields him some information regarding these maximum premiums $\overline{\pi}_i$ the employees are willing to pay. Based on this (possibly biased) information he can solve his optimization problem resulting in a premium $\hat{\pi}^*$. The problem is called implementable if the employer can design a game such that the outcome $\hat{\pi}^*$ of this game coincides with the optimal outcome π^*. The interpretation is that the employer gets perfectly informed on the maximum premiums when applying this specific game, since the outcome $\hat{\pi}^*$ of this game coincides with the outcome π^* that the employer would have chosen if he knew the maximum premiums beforehand. Summarizing, the employer's decision problem is implementable if he can construct a mechanism for which the outcomes coincide with the outcomes that are optimal from his point of view.

Unfortunately for the employer, it is shown that under rather weak conditions the employer's decision problem is not implementable. This means that there exists no mechanism, how clever or complex it may be designed, that provides the employer with the information he needs.

The paper is organized as follows. Section 2 introduces the model. Section 3 then states the Bayesian implementation framework. Furthermore, it is shown that the employer's decision problem is not Bayesian implementable. Finally, Section 4 concludes.

2 The Model

We start with explaining the model in the context of pension funds, that is, insurances that guarantee employees of a sufficiently high income after retirement. Then we show that by changing the meaning of some of the variables, this model also applies for disability insurances that guarantee employees of a sufficiently high income once they become disabled.

Let $N = \{1, 2, \ldots, n\}$ be the set of employees and denote the employer by E. The employer and the employees are assumed to be risk averse expected utility maximizers with utility functions u_i and u_E, respectively. Next, let $t = 1$ denote the starting year of the pension fund. Define for each employee $i \in N$

N_i the remaining lifetime of employee i;

P_i the year in which employee i plans to retire;

w_{it} the wage of employee i in year t, $w_{it} \geq 0$;

y_{it} the pension provided by the government to employee i in year t, $y_{it} \geq 0$;

E_{it} the additional pension payments in year t for which employee i can take an insurance.

It is assumed that both death and retirement occur at the end of a year. Furthermore, we introduce

τ tax rate on income, $t \in (0, 1)$;

g the social security premiums that have to be paid by the employer, expressed as a percentage of the employees' wages;

δ discount rate, $\delta \in [0, 1]$.

Note that the remaining lifetime N_i of employee i is a random variable and that P_i and E_{it} may be random variables[1]. Also note that wages and pension payments may vary over the years. Furthermore, let T be the year for which all employees have deceased with probability one. So, the time span that is considered covers the period from $t = 1$ to $t = T$. Next, we assume that pension/insurance premiums can be expressed as a percentage of the employee's wage w_{it}. Moreover, since it is illegal - at least in The Netherlands - for the employer to discriminate between his employees with respect to the insurance premium, each employee that participates in the pension fund pays the same percentage of his wage as insurance premium. Insurance companies operating on the market, however, are allowed to discriminate between the premiums that individuals have to pay for an insurance.

Now suppose that the employer organizes a pension fund for his employees and that the premium equals a percentage $\pi \in (0, 1)$ of the wage. Then employee $i \in N$ has three different actions at his disposal.

First, he can decide to take no insurance at all. Then his payoff in year t equals $(1-\tau)w_{it}$ if employee i is still working in year t, $(1-\tau)y_{it}$ if employee i is retired in year t, and zero if employee i is no longer alive in year t. Thus, employee i's discounted payoff over the period of T years equals

$$\sum_{t=1}^{\min(P_i,N_i)} \delta^t(1-\tau)w_{it} \; + \sum_{t=P_i+1}^{N_i} \delta^t(1-\tau)y_{it}, \tag{1}$$

where the latter sum equals zero if $P_i + 1 > N_i$.

Second, employee i can decide to insure his income after retirement at an existing insurance company. Then his payoff in year t equals $(1-\tau)w_{it} - \pi_i^m w_{it}$ if employee i is still working in year t. Note that the premium $\pi_i^m w_{it}$ employee i pays in year t is a percentage π_i^m of his wage and that it may depend on i. His payoff equals $(1-\tau)(y_{it} + E_{it})$ if employee i is already retired in year t. Recall that E_{it} equals the payment that is provided by the insurance and that it may vary over the years. Finally, his payoff equals zero if employee i is no longer alive in year t. So, employee i's discounted payoff over T years equals

$$\sum_{t=1}^{\min(P_i,N_i)} \delta^t((1-\tau)w_{it} - \pi_i^m w_{it}) \; + \sum_{t=P_i+1}^{N_i} \delta^t(1-\tau)(y_{it} + E_{it}). \tag{2}$$

The third and final possibility for employee i is to participate in the employer's pension fund. Then his payoff equals $(1-\tau)(w_{it} - \pi w_{it})$ if employee i is working in year t. Note

[1] We assume that all random variables are measurable functions with finite expectations.

that the premium πw_{it} employee i pays in year t may be deducted from his gross wage, which decreases the total amount of tax payments. His payoff equals $(1-\tau)(y_{it}+E_{it})$ if employee i has already retired in year t, and his payoff equals zero if employee i is no longer alive in year t. So, in this case employee i's discounted payoff over T years equals

$$\sum_{t=1}^{\min(P_i,N_i)} \delta^t(1-\tau)(w_{it}-\pi_{it}^w) \; + \sum_{t=P_i+1}^{N_i} \delta^t(1-\tau)(y_{it}+E_{it}). \tag{3}$$

Out of the three options considered above, employee i chooses the one that maximizes his expected utility. Moreover, taking the premiums charged by the existing insurance companies as given, employee i can determine the maximum premium π he is willing to pay at most for participation in the pension fund. Let $\overline{\pi}_i$ denote employee i's maximum premium. Then without loss of generality we may assume that $\overline{\pi}_1 \geq \overline{\pi}_2 \geq \ldots \geq \overline{\pi}_n$.

Given these maximum premiums $\overline{\pi}_i$ we can consider the employer's decision problem. First, of course, the employer has to decide whether or not he starts a pension fund for his employees. If he decides to do so, then he has to determine the premium that the employees have to pay for participating in the pension fund. Furthermore, he has to decide how much he contributes himself to the pension fund's capital every year. This contribution in year t is denoted by $c_t \in \mathbb{R}$. An incentive for the employer to contribute to the pension fund is the following. The contributions c_t enable the employer to lower the premium, possibly resulting in more employees participating in the pension fund, which yields that less social premiums have to be paid. Hence, the benefits of the employer may increase.

Suppose the employer does not organize a pension fund for his employees. In that case his payoff in year t equals $-\sum_{i=1}^{n}\sum_{t=1}^{\min(P_i,N_i)}(1+g)w_{it}$, that is, the wages he has to pay to his employees plus the social premiums.

Next, suppose that he does organize a pension fund. Then, given a premium π, employee i participates in the pension fund if $\pi \leq \overline{\pi}_i$, that is, if the premium is less than or equal to the maximum premium he is willing to pay for the insurance. So in case of indifference we assume that he participates. Then the number $m(\pi)$ of employees that participate equals $m(\pi) = |\{i \in N | \pi \leq \overline{\pi}_i\}|$. Moreover, since we assumed that $\overline{\pi}_1 \geq \overline{\pi}_2 \geq \ldots \geq \overline{\pi}_n$ it holds that $\overline{\pi}_i \geq \pi$ for $i = 1, 2, \ldots, m(\pi)$ and $\overline{\pi}_i < \pi$ for $i = m(\pi)+1, \ldots, n$. The employer's payoff in year t with respect to a working employee i equals $-(1+g)(w_{it} - \pi w_{it})$ if employee i participates in the pension fund, and it equals $-(1+g)w_{it}$ if he does not. Note that the premium may be deducted from the employee's gross wage, resulting in a decrease in social premiums that have to be paid. With respect to a retired employee i, the employer's payoff in year t equals $-E_{it}$ if employee i participates in the pension fund, and it equals zero if he does not. Furthermore, the employer contributes in year t an amount c_t to the pension fund's capital. So, the discounted payoff over the period of T years equals

$$
\begin{aligned}
&-\sum_{i=1}^{n} \sum_{t=1}^{\min(P_i,N_i)} \delta^t(1+g)w_{it} \;+\; \sum_{i=1}^{m(\pi)} \sum_{t=1}^{\min(P_i,N_i)} \delta^t g\pi w_{it} \\
&-\sum_{i=1}^{m(\pi)} \sum_{t=P_i+1}^{N_i} \delta^t E_{it} \;-\; \sum_{t=1}^{T} \delta^t c_t.
\end{aligned}
\tag{4}
$$

There are, however, solvability constraints on the pension fund's capital that have to be satisfied. For each year t the probability that the total amount of money claimed thus far exceeds the available capital must be sufficiently small. The total discounted claims up to year t equal

$$
\sum_{i=1}^{m(\pi)} \sum_{k=P_i+1}^{\min(t,N_i)} \delta^k E_{ik},
$$

and the available capital up to year t equals

$$
\sum_{i=1}^{m(\pi)} \sum_{k=1}^{\min(t,P_i,N_i)} \delta^k \pi w_{ik} \;+\; \sum_{k=1}^{t} \delta^k c_k.
$$

Note that both the claims and the available capital are random variables. The solvability constraints then read as follows,

$$
\mathbb{P}\left(\sum_{i=1}^{m(\pi)} \sum_{k=P_i+1}^{\min(t,N_i)} \delta^k E_{ik} \;>\; \sum_{i=1}^{m(\pi)} \sum_{k=1}^{\min(t,N_i,P_i)} \delta^k \pi w_{ik} \;+\; \sum_{k=1}^{t} \delta^k c_k \right) \;\leq\; \alpha,
\tag{5}
$$

for $t = 1, 2, \ldots, T$ and α sufficiently small, e.g., $\alpha = 0.05$.

Summarizing, the employer chooses π and $c_t(t = 1, 2, \ldots, T)$ such that his expected utility is maximal given that the solvability restrictions of (5) are satisfied. So, the employer's decision problem can be stated as follows,

$$
\begin{aligned}
\max_{\pi,c_t} \quad & E\left(u_E\left(-\sum_{i=1}^{n} \sum_{t=1}^{\min(P_i,N_i)} \delta^t(1+g)w_{it} \;+\; \sum_{i=1}^{m(\pi)} \sum_{t=1}^{\min(P_i,N_i)} \delta^t g\pi w_{it} \right.\right. \\
& \left.\left. -\sum_{i=1}^{m(\pi)} \sum_{t=P_i+1}^{N_i} \delta^t E_{it} \;-\; \sum_{t=1}^{T} \delta^t c_t \right)\right) \\
\text{s.t.:} \quad & |\{i \in N | \pi \leq \overline{\pi}_i\}| = m(\pi) \\
& \mathbb{P}\left(\sum_{i=1}^{m(\pi)} \sum_{k=P_i}^{\min(t,N_i)} \delta^k E_{ik} \;>\; \sum_{i=1}^{m(\pi)} \sum_{k=1}^{\min(t,N_i,P_i)} \delta^k \pi w_{ik} \;+\; \sum_{k=1}^{t} \delta^k c_k \right) \;\leq\; \alpha \\
& \text{for } t = 1, 2, \ldots, T.
\end{aligned}
\tag{6}
$$

Note that if the resulting expected utility is less than the utility level corresponding to organizing no pension fond, then the employer does, of course, not establish a pension fund for his employees.

Next, let us take a closer look at the employer's decision problem. To start with the solvability constraints, note that probability distribution functions are continuous from the right. So, given a premium π there exists a contribution $c_1(\pi)$ of the employer such that the solvability constraint for $t = 1$ is satisfied in equality, that is,

$$\mathbb{P}\left(\sum_{i=1}^{m(\pi)} \sum_{k=P_i+1}^{\min(1,N_i)} \delta^k E_{ik} \;>\; \sum_{i=1}^{m(\pi)} \sum_{k=1}^{\min(1,N_i,P_i)} \delta^k \pi w_{ik} \;+\; \delta c_1(\pi)\right) \;=\; \alpha,$$

and

$$\mathbb{P}\left(\sum_{i=1}^{m(\pi)} \sum_{k=P_i+1}^{\min(1,N_i)} \delta^k E_{ik} \;>\; \sum_{i=1}^{m(\pi)} \sum_{k=1}^{\min(1,N_i,P_i)} \delta^k \pi w_{ik} \;+\; \delta x\right) \;>\; \alpha,$$

for all $x < c_1(\pi)$. Thus, if π is the premium then $c_1(\pi)$ is the optimal contribution in year $t = 1$ of the employer to the pension fund's capital. In a similar way, one can calculate the optimal contributions $c_t(\pi)$ for the years $t = 2, 3, \ldots, T$. So, the employer's optimization problem can be written as

$$\begin{aligned} \max_{\pi} \quad & E\left(u_E\left(-\sum_{i=1}^{n} \sum_{t=1}^{\min(P_i,N_i)} \delta^t (1+g) w_{it} \;+\; \sum_{i=1}^{m(\pi)} \sum_{t=1}^{\min(P_i,N_i)} \delta^t g \pi w_{it}\right.\right. \\ & \left.\left. -\sum_{i=1}^{m(\pi)} \sum_{t=P_i+1}^{N_i} \delta^t E_{it} \;-\; \sum_{t=1}^{T} \delta^t c_t(\pi) - \right)\right) \\ \text{s.t.:} \quad & |\{i \in N | \pi \leq \overline{\pi}_i\}| = m(\pi). \end{aligned} \tag{7}$$

Next, we show that in the optimum the premium π can be chosen equal to the maximum premium $\overline{\pi}_i$ of some employee i. For this purpose, let π be a premium such that $\pi \neq \overline{\pi}_i$ for all $i = 1, 2, \ldots, n$. Furthermore, let $m = m(\pi)$ be the number of insured employees if the premium equals π. The employer can (weakly) increase his expected utility by setting the premium equal to $\overline{\pi}_m$, the premium that employee m is willing to pay at most for the insurance. To see this, first note that $\overline{\pi}^m > \pi$. This implies that the payoff $-\sum_{i=1}^{m} \sum_{t=1}^{\min(P_i,N_i)} \delta^t \pi w_{it}$ stochastically dominates[2] the payoff $-\sum_{i=1}^{m} \sum_{t=1}^{\min(P_i,N_i)} \delta^t \overline{\pi}_m w_{it}$. Hence,

$$\mathbb{P}\left(\sum_{i=1}^{m} \sum_{t=P_i+1}^{N_i} \delta^t E_{it} - \sum_{i=1}^{m} \sum_{t=1}^{\min(P_i,N_i)} \delta^t \pi w_{it} \;\leq\; x\right) \;\leq$$
$$\mathbb{P}\left(\sum_{i=1}^{m} \sum_{t=P_i+1}^{N_i} \delta^t E_{it} - \sum_{i=1}^{m} \sum_{t=1}^{\min(P_i,N_i)} \delta^t \overline{\pi}_m w_{it} \;\leq\; x\right)$$

[2]Let X and Y be random variables. X stochastically dominates Y if $\mathbb{P}(X \leq x) \leq \mathbb{P}(Y \leq x)$ for all $x \in \mathbb{R}$.

for all $x \in \mathbb{R}$, which implies that $\sum_{t=1}^{T} \delta^t c_t(\overline{\pi}_m) \leq \sum_{t=1}^{T} \delta^t c_t(\pi)$. Then the fact that payoff $\sum_{i=1}^{m} \sum_{t=1}^{\min(P_i,N_i)} \delta^t g \overline{\pi}_m w_{it}$ stochastically dominates the payoff $\sum_{i=1}^{m} \sum_{t=1}^{\min(P_i,N_i)} \delta^t g \pi w_{it}$ implies that

$$
\begin{aligned}
&E\left(u_E\left(-\sum_{i=1}^{n} \sum_{t=1}^{\min(P_i,N_i)} \delta^t (1+g) w_{it} \;+\; \sum_{i=1}^{m(\pi)} \sum_{t=1}^{\min(P_i,N_i)} \delta^t g \overline{\pi}_m w_{it} \right.\right.\\
&\qquad \left.\left. -\sum_{i=1}^{m(\pi)} \sum_{t=P_i+1}^{N_i} \delta^t E_{it} \;-\; \sum_{t=1}^{T} \delta^t c_t(\overline{\pi}_m) - \right)\right) \;\geq \\
&E\left(u_E\left(-\sum_{i=1}^{n} \sum_{t=1}^{\min(P_i,N_i)} \delta^t (1+g) w_{it} \;+\; \sum_{i=1}^{m(\pi)} \sum_{t=1}^{\min(P_i,N_i)} \delta^t g \pi w_{it} \right.\right.\\
&\qquad \left.\left. -\sum_{i=1}^{m(\pi)} \sum_{t=P_i+1}^{N_i} \delta^t E_{it} \;-\; \sum_{t=1}^{T} \delta^t c_t(\pi) - \right)\right).
\end{aligned}
$$

Since the payoff corresponding to a premium π can be (weakly) improved by setting the premium equal to $\overline{\pi}_{m(\pi)}$, an optimal premium can be found among the maximum premiums $\overline{\pi}_i$, $i = 1, 2, \ldots, n$. So, instead of deciding what premium to charge he can decide on how many employees he wants to insure. If he wants to insure exactly m employees, he sets the premium equal to $\overline{\pi}_m$. Note, however, that some values of m are excluded if several maximum premiums coincide. For example, if $\overline{\pi}_m = \overline{\pi}_{m+1}$ then it is not possible for the employer to insure exactly m employees. For setting the premium equal to $\overline{\pi}_m$ yields $m+1$ insured employees. Formally, this means that the employer can only choose among those m for which $\overline{\pi}_m > \overline{\pi}_{m+1}$. Let D denote this domain, i.e.,

$$
D = \{m \in \{1, 2, \ldots, n\} | \overline{\pi}_m > \overline{\pi}_{m+1}\} \cup \{0\}.
$$

Note that we have also included $m = 0$, so that organizing no pension fund is also incorporated as a possibility. Exchanging the decision variable π for m in the employer's decision problem then yields

$$
\begin{aligned}
\max_{m \in D}\; &E\left(u_E\left(-\sum_{i=1}^{n} \sum_{t=1}^{\min(P_i,N_i)} \delta^t (1+g) w_{it} \;+\; \sum_{i=1}^{m} \sum_{t=1}^{\min(P_i,N_i)} \delta^t g \overline{\pi}_m w_{it} \right.\right.\\
&\qquad \left.\left. -\sum_{i=1}^{m} \sum_{t=P_i+1}^{N_i} \delta^t E_{it} \;-\; \sum_{t=1}^{T} \delta^t c_t(\overline{\pi}_m) - \right)\right)
\end{aligned} \tag{8}
$$

Remark that the employer's constrained optimization problem has now been reduced to an unconstrained optimization problem over a finite number of possibilities.

The problem considered thus far is concerned with the insurance of an (additional) income after retirement. Now, by changing the meanings of the variables N_i and P_i, this model also describes the insurance of an (additional) income when an employer becomes disabled. For this, define for each employee i

- N_i the year in which the insurance payments to employee i end. This, for example, happens when employee i dies or when he reaches the retirement age, in which case he receives a pension from the government.;
- P_i the year in which employee i becomes disabled;
- w_{it} the wage of employee i in year t, $w_{it} \geq 0$;
- y_{it} the payments provided by the government to employee i in year t, $y_{it} \geq 0$;
- E_{it} the additional income in year t for which employee i can take an insurance;

So, an employee pays premiums for the insurance from $t = 1$ to $t = \min(P_i, N_i)$, and he receives payments from this insurance from $t = P_i + 1$ till $t = N_i$. Note that these are the same time periods that occur in the model on pensions. Thus the problem formulated in (8) also applies if the insurance provided by the employer concerns the disability of an employee.

Both insurance problems assume that the employer is perfectly informed on the maximum premiums $\overline{\pi}_i$ of the employees. Given these maximum premiums he can determine the optimal solution m^* of (8), and subsequently he can offer his employees to participate in the pension fund for a premium $\overline{\pi}_{m^*}$. In reality, however, it is unlikely that an employer is completely informed on the maximum premiums $\overline{\pi}_i$ his employees are willing to pay for an insurance. Indeed, this maximum premium is private information for the employee and not known to his colleagues or his employer.

Since the employer has incomplete knowledge of the maximum premiums $\overline{\pi}_i$ of his employees, he is not able to solve the optimization problem formulated in (8). To get information on these maximum premiums, the employer could approach each of his employees and ask them to reveal their maximum premiums. Subseqeuntly, he could determine the optimal solution as described above and announce the premium to his employees. Obtaining information in this way, however, has the drawback that employees have incentives to misreport their maximum premiums. More precisely, employees have a tendency to reveal lower maximum premiums than the true ones. As a result, the inaccurate information used by the employer results in a premium that is not optimal for him. Arises the question whether there is some other way for the employer to obtain the information he needs. This question is dealt with in the next section.

3 The Implementation Problem

Acquiring the correct information in order to determine the optimal premium from the employer's point of view is known in the literature as an implementation problem. To be more specific, since employees only have information on their own maximum premium $\overline{\pi}_i$ and not on the maximum premiums of the other employees, it is a Bayesian implementation problem. Before we take a more general approach to Bayesian implementation let us elaborate on the situation described at the end of the previous section: the situation where each employee was asked to report his maximum premium.

Let $N = \{1, 2, \ldots, n\}$ denote the set of employees. The maximum premium $\overline{\pi}_i$ employee i is willing to pay is determined by his utility function u_i. So the fact that the employer does not know employee i's maximum premium implies that he also does not know his utility function. Furthermore, employee i only knows his own utility function and not the utility functions of all the other employees. Let $u = (u_1, u_2, \ldots, u_n)$ denote the vector of utility functions. Such a vector u is called the state of the world. Recall that the employer has no knowledge about the state u of utility functions and that employee i only has partial knowledge of the state u, namely, his own utility function u_i. Furthermore, since employees were assumed to be risk averse expected utility maximizers the set of all possible utility functions for employee i equals

$$U_i = \{u_i : \mathbb{R} \to \mathbb{R} | u_i \text{ is strictly increasing and concave}\},$$

and the set of all possible state of the world equals $U = U_1 \times U_2 \times \ldots \times U_n$.

To get some information on the true state of the world the employer asks his employees to reveal their utility functions. Suppose that $u = (u_1, u_2, \ldots, u_n)$ denotes the true state of the world and that $\hat{u} = (\hat{u}_1, \hat{u}_2, \ldots, \hat{u}_n)$ denotes the reported state of the world. Given the reported utility functions $\hat{u}_i$, the employer can use expression (3) to determine the maximum premium $\overline{\pi}(\hat{u}_i)$ each employee is willing to pay for the insurance. Subsequently, he uses these maximum premiums to solve his optimization problem formulated in (8), yielding an optimal premium $\pi^*(\hat{u})$ that depends on the reported utility functions $\hat{u} = (\hat{u}_1, \hat{u}_2, \ldots, \hat{u}_n)$. Then the employer rules that employees have to participate in the pension fund if the premium $\pi^*(\hat{u})$ charged by the employer is less than or equal to the reported maximum premium $\overline{\pi}(\hat{u}_i)$. Employees are not allowed to participate if the premium $\pi^*(\hat{u})$ exceeds the reported maximum premium. Moreover, the employees know beforehand that this rule applies. So, they can take this into account when deciding which utility function to reveal. At first sight this rule might seem a bit restrictive compared to the more realistic case where the employees may decide themselves whether or not they want to participate in the pension fund when the premium equals $\pi^*(\hat{u})$. We will show later on though, that for our main result it does not matter which of the two rules apply.

Given the premium $\pi^*(\hat{u})$ employee i participates in the pension fund if $\overline{\pi}(\hat{u}_i) \geq \pi^*(\hat{u})$, yielding the payoff (cf. (3))

$$\sum_{t=1}^{\min(P_i,N_i)} \delta^t(1-\tau)(w_{it} - \pi^*(u)w_i t) \; + \sum_{t=P_i+1}^{N_i} \delta^t(1-\tau)(y_{it} + E_{it}).$$

On the other hand, if $\overline{\pi}(\hat{u}_i) < \pi^*(\hat{u})$ employee i does not participate. His payoff then equals

$$\sum_{t=1}^{\min(P_i,N_i)} \delta^t(1-\tau)w_{it} \; + \sum_{t=P_i+1}^{N_i} \delta^t(1-\tau)y_{it},$$

or

$$\sum_{t=1}^{\min(P_i,N_i)} \delta^t((1-\tau)w_{it} - \pi_i^m w_i t) \; + \sum_{t=P_i+1}^{N_i} \delta^t(1-\tau)(y_{it} + E_{it}),$$

depending on which of the two outside options employee i prefers most, that is, no insurance or insurance with an existing insurance company (cf. (1) and (2)).

Let $h_i^*(\hat{u})$ denote the payoff to employee i when $\hat{u}$ is the reported state of the world and let $h^*(\hat{u}) = (h_1^*(\hat{u}), h_2^*(\hat{u}), \ldots, h_n^*(\hat{u}))$. The expected utility for employee i then equals $E(u_i(h_i^*(\hat{u})))$. Note that the payoff $h_i^*(\hat{u})$ is evaluated by his true utility function u_i. Since $h_i^*(\hat{u})$ depends on $\hat{u}_i$ employee i can influence the outcome by reporting another utility function. Which utility function he reports depends on the strategy he plays.

A strategy σ_i for employee i indicates which utility function is reported given the true utility function. Formally, this means that σ_i is a map assigning to each utility function $u_i \in U_i$ another utility function $\hat{u}_i \in U_i$, with the interpretation that if u_i represents employee i's true utility function then he reports the utility function $\hat{u}_i$. For example, the strategy that always reports the true utility function is represented by $\sigma_i(u_i) = u_i$ for all $u_i \in U_i$.

Given the strategies $\sigma = (\sigma_1, \sigma_2, \ldots, \sigma_n)$ of all employees and the true state of the world u the payoff to employee i equals $h_i^*(\sigma(u))$, where $\sigma(u) = (\sigma_1(u_1), \sigma_2(u_2), \ldots, \sigma_n(u_n))$.

The corresponding expected utility then equals $E(u_i(h_i^*(\sigma(u))))$. Determining this expected utility, however, is impossible considering the fact that employee i does not completely know the state u. In fact, he only knows his own utility function u_i. To circumvent this problem, each employee i has beliefs about the true state u. These beliefs are represented by a prior probability measure $\mu_i(u_i)$ on $U_{-i} = \Pi_{j \neq i} U_j$ and may depend on the true utility function u_i of employee i. The expected utility of employee i is then given by

$$\int_{U_{-i}} E(u_i(h_i^*(\sigma(u_i, u_{-i})))) d\mu_i(u_i).$$

Furthermore, given that σ_{-i} describe the strategies of the other employees, employee i (weakly) prefers a strategy σ_i to a strategy $\hat{\sigma}_i$ if the expected utility of σ_i is greater than or equal to the expected utility of $\hat{\sigma}_i$, that is,

$$\int_{U_{-i}} E(u_i(h_i^*(\sigma_i(u_i), \sigma_{-i}(u_{-i})))) d\mu_i(u_i) \geq \int_{U_{-i}} E(u_i(h_i^*(\hat{\sigma}_i(u_i), \sigma_{-i}(u_{-i})))) d\mu_i(u_i),$$

where $\sigma_{-i}(u_{-i}) = (\sigma_j(u_j))_{j \neq i}$.

A strategyvector σ is called a Bayesian equilibrium if for each employee i and each possible true utility function u_i the strategy σ_i maximizes his expected utility given the strategies σ_{-i} of the other employees. Mathematically, this means that for each $i \in N$, each $u_i \in U_i$, and each $\hat{\sigma}_i$ it holds that

$$\int_{U_{-i}} E(u_i(h_i^*(\sigma_i(u_i), \sigma_{-i}(u_{-i})))) d\mu_i(u_i) \geq \int_{U_{-i}} E(u_i(h_i^*(\hat{\sigma}_i(u_i), \sigma_{-i}(u_{-i})))) d\mu_i(u_i). \tag{9}$$

The payoff function h^* is called incentive compatible if reporting the true utility function is a Bayesian equilibrium. Substituting in (9) that reporting the truth is the strategy $\sigma(u) = u$ for all $u \in U$ and that $\hat{\sigma}_i(u_i) = \hat{u}_i$ for some $\hat{u}_i \in U_i$ implies that h^* is incentive compatible if

$$\int_{U_{-i}} E(u_i(h_i^*(u_i, u_{-i}))) d\mu_i(u_i) \geq \int_{U_{-i}} E(u_i(h_i^*(\hat{u}_i, u_{-i}))) d\mu_i(u_i) \tag{10}$$

holds for all $i \in N$ and all $u_i, \hat{u}_i \in U_i$.

Summarizing, if h^* is incentive compatible then revealing the true utility functions is a Bayesian equilibrium. This means from an equilibrium point of view that the employer becomes correctly informed on the true utility functions and, consequently, that he charges the optimal insurance premium that yields him maximal benefits. The next theorem, however, shows that under some conditions the payoff function h^* does not satisfy incentive compatibility. Hence, the employer can not expect to get perfect information on the true utility functions.

To specify the conditions take $i \in N$ and $u_i \in U_i$. Let $\Pi_{-i} \subset \mathbb{R}^{n-1}$ be the set of all possible maximum premiums $\overline{\pi}(u_{-i}) = (\overline{\pi}(u_j))_{j \neq i}$. Thus,

$$\Pi_{-i} = \{\overline{\pi}(u_{-i}) | u_{-i} \in U_{-i}\}.$$

Next, let B be a Borel set in Π_{-i} and define the set $V(B) \subset U_{-i}$ by

$$V(B) = \{u_{-i} \in U_{-i} | \overline{\pi}(u_{-i} \in B\},$$

as the set of all utility functions u_{-i} of the other employees such that the corresponding vector of maximum premiums belongs to B.

Theorem 3.1 Let λ denote the Lebesgue measure on $\mathbb{R}^{n-1}$. If there exists an employee $i \in N$ and a utility function $u_i \in U_i$ with a prior probability measure $\mu_i(u_i)$ such that for every Borel set $B \subset \Pi_{-i}$ it holds that

$$\lambda(B) > 0 \text{ if and only if } \mu_i(u_i)(V(B)) > 0,$$

then h^* is not incentive compatible.

PROOF: See Appendix A.

The condition is satisfied for sure if there exists an employee i and a utility function u_i such that the density function corresponding to this employee's prior probability measure $\mu_i(u_i)$ is strictly positive.

Since the payoff function h^* is not incentive compatible, the employer cannot get the information he needs by just asking his employees. This is, however, only one way to get information. The employer could, of course, design other methods to receive information on the maximum premiums of his employees. These other methods are described by so-called mechanisms.

A mechanism is a pair (A, α), where $A = A_1 \times A_2 \times \ldots \times A_n$ with A_i the action space of employee i and α the payoff function, assigning to each action $a \in A$ a (stochastic) payoff to each employee. For example, in the previously described mechanism where each employee reports his utility function, an action for employee i consists of reporting a utility function.

This implies that $A_i = U_i$ for all $i \in N$ and that $A = U$. Furthermore, the payoff function α is described by h^*, i.e., $\alpha(a) = h^*(a)$ for all $a \in A$.

Similar to the mechanism (U, h^*) we can define strategies for employee i in a mechanism (A, α). A strategy for employee i is a map $\sigma_i : U_i \to A_i$ with the interpretation that employee i takes action $\sigma_i(u_i) \in A_i$ if u_i is his true utility function.

Given the strategies σ of all employees, the payoff to employee i equals $\alpha_i(\sigma(u))$ if u is the vector of true utility functions. The expected payoff to employee i then equals

$$\int_{U_{-i}} E(u_i(\alpha_i(\sigma(u_i, u_{-i}))))d\mu_i(u_i).$$

Furthermore, employee i (weakly) prefers the strategy σ_i to the strategy $\hat{\sigma}_i$ if the expected utility of σ_i is greater than or equal to the expected utility of $\hat{\sigma}_i$. Formally, this means that

$$\begin{aligned}\int_{U_{-i}} E(u_i(\alpha_i(\sigma_i(u_i), \sigma_{-i}(u_{-i}))))d\mu_i(u_i) \;\geq& \\ \int_{U_{-i}} E(u_i(\alpha_i(\hat{\sigma}_i(u_i), \sigma_{-i}(u_{-i}))))d\mu_i(u_i),&\end{aligned}$$

A strategyvector σ is then called a Bayesian equilibrium if for each employee i and each possible true utility function u_i the strategy σ_i maximizes his expected utility given the strategies σ_{-i} of the remaining employees. Mathematically, σ is a Bayesian equilibrium if for each $i \in N$, each $u_i \in U_i$ and each $\hat{\sigma}_i$ it holds that

$$\begin{aligned}\int_{U_{-i}} E(u_i(\alpha_i(\sigma_i(u_i), \sigma_{-i}(u_{-i}))))d\mu_i(u_i) \;\geq& \\ \int_{U_{-i}} E(u_i(\alpha_i(\hat{\sigma}_i(u_i), \sigma_{-i}(u_{-i}))))d\mu_i(u_i),&\end{aligned} \tag{11}$$

Now, consider again the problem from the employer's point of view. If $u \in U$ are the true utility functions then he would like the payoffs to the employees to be equal to $h^*(u)$, since this corresponds to maximal benefits for the employer. To be sure that his benefits are maximal for each possible state u, the employer has to design the mechanism (A, α) in such a way that every Bayesian equilibrium σ yields him the payoff $\alpha(\sigma(u)) = h^*(u)$ for all $u \in U$. If he can indeed construct such a mechanism then the payoff function h^* is called Bayesian implementable.

Similar to Bayesian implementation with a finite number of states of the world (see *Jackson (1991)*) one can easily show that incentive compatibility as formulated in (10) is a necessary condition or h^* to be Bayesian implementable. To see this, suppose that h^* is Bayesian implementable but not incentive compatible. Then some employee i is better off in the revelation mechanism (U, h^*) by pretending that some $\hat{u}_i$ is his true utility function instead of his true utility function u_i. But since each Bayesian equilibrium σ in the mechanism (A, α) yields him the payoff $\alpha_i(\sigma(u)) = h_i^*(u)$ he could improve his payoff in this same mechanism (A, α) by acting as if $\hat{u}_i$ is his utility function instead of u_i, that is, take the action $\sigma_i(\hat{u}_i)$ instead of $\sigma_i(u_i)$. Obviously, this contradicts the fact that σ is a Bayesian equilibrium. Hence, the payoff function h^* must satisfy incentive compatibility.

Theorem 3.1 showed that under rather weak conditions, the payoff function h^* does not satisfy the incentive compatibility property. This implies that h^* is also not Bayesian implementable. Thus there exists no mechanism, how complex or clever it may be designed, that gives the employer the correct information regarding the utility functions of his employees.

4 Remarks

Theorem 3.1 tells us that under some conditions the employer is not able to perfectly inform himself on the maximum premiums of his employees. Recall that the payoff function desired by the employer was chosen in such a way that employees can only participate in the pension fund if the maximum premium they report does not exceed the premium set by the employer. So, in fact, the employer decided whether an employee got the insurance or not. As already mentioned, a natural (and maybe more realistic) adjustment of this payoff function is to let the employees decide whether they want to participate or not. So, given the information revealed by the employees, the employer sets an insurance premium and each employee decides individually if he participates or not. This payoff function, however, does also not satisfy the incentive compatibility property. This can easily be seen as follows. Suppose that it did satisfy incentive compatibility. Then revealing the true utility function would be a Bayesian equilibrium. But if each employee reveals his true utility function then it does not matter whether the employer decides on the participation of an employee in the pension fund, or the employee himself decides this. So if the adjusted payoff function satisfies incentive compatibility, then so does the original payoff function. Since the latter is not true, the adjusted payoff function can not be incentive compatible.

In proving the impossibility theorem we have put no restrictions on the set of possible utility functions U_i. So U_i contains any strictly increasing and concave utility function. This guarantees that for any 'reasonable' insurance premium there exists a utility function such that the corresponding maximum premium equals the aforementioned insurance premium. In this context, reasonable means that the premium should be at least the expected loss of income. This is due to the risk averse behavior of the employees. Furthermore, the premium should not exceed the maximum amount that theoretically can be received from the insurance. Taking this notion of reasonability into account, it is not difficult to see that Theorem 3.1 still holds if the set of utility functions U_i is restricted in such a way, that for every 'reasonable' insurance premium there still exists a utility function in U_i for which the corresponding maximum premium equals this insurance premium. An example of such a restricted U_i is $U_i = \{u_i(t) = -e^{-a_i t}, t \in \mathbb{R} | a_i > 0\}$. An incidental benefit in this case is that the beliefs of employee i on the utility functions of the remaining employees can be described by a probability distribution on a_{-i}, i.e., the parameters that determine the utility functions of the other employees. Furthermore, the condition formulated in Theorem 3.1 holds for any continuous probability distribution on a_{-i} with strictly positive density function.

The overall consequence of Theorem 3.1 is that there is no need for the employer to organize a (complex) mechanism to acquire information when setting up a pension fund for his employees, since all the information revealed by the employees cannot be rated at its true value. The alternatives left for the employer are either to get the information elsewhere (if possible) or to set an insurance premium independent of the information revealed by the employees. for the actual height of the premium the employer could use the market premium as an indication.

Finally, it should be noted that a mechanism (A, α) includes many ways of collecting information on the maximum premiums of the employees. For example, by introducing individual taxes the employer can discriminate between individual employees. Although such a mechanism is illegal in our model, the main result does not change if it would be allowed. So, even if the employer is allowed to discriminate between his employees, he can not get perfectly informed on the maximum premiums.

5 Appendix: Proof of Theorem 3.1

The outline of the proof is as follows. To show our main result we only consider deviations by employees to utility functions yielding a lower maximum premium. This turns out to be sufficient to show that h^* is not incentive compatible. The proof consists of five parts. In the first part of the proof we examine how the premium changes when an employee deviates from revealing his true utility function. The second part focuses on the cases for which deviating decreases the expected utility of an employee, while the third part considers the cases for which deviating increases his expected utility. In the fourth part of the proof we deduce a necessary condition on h^* to satisfy incentive compatibility. The fifth and final part then shows that this necessary condition cannot be satisfied.

Throughout this proof let the true utility function u_i of employee i be fixed and let $\overline{\pi}(u_i)$ denote the maximum premium corresponding to u_i. Furthermore, let $\pi^*(\hat{u})$ denote the optimal premium set by the employer when utility functions $\hat{u}$ are revealed. Finally, we make the following assumptions. First, all employees are identical. So, N_i, P_i, w_{it}, y_{it}, and E_{it} do not depend on i. Second, since $\{x \in \mathbb{R}^n | x_i = \overline{\pi}(u_i), \exists_{j \in N} : x_j = \overline{\pi}(u_i)\}$ is a set with Lebesgue measure zero, it follows that $\{\hat{u}_{-i} \in U_{-i} | \exists_{j \in N} : \overline{\pi}(\hat{u}_j) = \overline{\pi}(u_i)\}$ has probability measure zero with respect to $\mu_i(u_i)$. Hence, we can restrict our attention to

$$\hat{U} = \{\hat{u} \in U | \forall_{j,i \in N} : \overline{\pi}(\hat{u}_j) \neq \overline{\pi}(u_i)\}.$$

So, in every state of the world no two employees can have the same maximum premium. Third, we assume that if an employee deviates then he reports a utility function with a lower maximum premium than the one corresponding to his true utility function.

Let us start with examining how the premium π^* changes when employee i reveals $\hat{u}_i$ instead of his true utility function u_i. Note that by assumption $\overline{\pi}(u_i) > \overline{\pi}(\hat{u}_i)$. Consider $\hat{u}_{-i} \in \hat{U}_{-i}$ and let $\overline{\pi}_m(\tilde{u})$ denote the m-highest maximum premium when $\tilde{u}$ are the reported utility functions, i.e., $\overline{\pi}_1(\tilde{u}) \geq \overline{\pi}_2(\tilde{u}) \geq \ldots \geq \overline{\pi}_n(\tilde{u})$. The employer's optimization problem formulated in (8) can be written as

$$\max_{m \in D(\tilde{u})} E\left(u_E\left(-W + \sum_{i=1}^{m}(\overline{\pi}_m(\tilde{u})W_i - E_i) - c(\overline{\pi}_m(\tilde{u}))\right)\right), \qquad (12)$$

where

$$
\begin{aligned}
W &= \sum_{i=1}^{n} \sum_{t=1}^{\min(P_i,N_i)} \delta^t (1+g) w_{it}, \\
W_i &= \sum_{t=1}^{\min(P_i,N_i)} \delta^t (1+g) w_{it}, \text{ for } i = 1,2,\ldots,n \\
E_i &= \sum_{t=P_i+1}^{N_i} \delta^t E_{it}, \text{ for } i = 1,2,\ldots,n \\
c(\overline{\pi}_m(\tilde{u})) &= \sum_{t=1}^{T} \delta^t c_t(\overline{\pi}_m(\tilde{u})) \text{ and} \\
D(\tilde{u}) &= \{m \in \{1,2,\ldots,n\} | \overline{\pi}_m(\tilde{u}) > \overline{\pi}_{m+1}(\tilde{u})\} \cup \{0\}.
\end{aligned}
$$

To see how the premium π^* changes when employee i reveals the utility function $\hat{u}_i$ instead of u_i, we distinguish three cases, namely, $\overline{\pi}(u_i) < \pi^*(u_i, \hat{u}_{-i})$, $\overline{\pi}(u_i) > \pi^*(u_i, \hat{u}_{-i})$, and $\overline{\pi}(u_i) = \pi^*(u_i, \hat{u}_{-i})$. These cases are graphically presented in Figure 1 on page 27 . In each graph the height of a bar represents the maximum premium that is reported by an employee. Note that the employees are placed in decreasing order with respect to the reported maximum premium.

Case 1: In this situation we have that $\overline{\pi}(u_i) < \pi^*(u_i, \hat{u}_{-i})$. So, employee i does not get an insurance from the employer. Now, suppose that employee i reports a utility function $\hat{u}_i$ with $\overline{\pi}(\hat{u}_i) < \overline{\pi}(u_i)$. Note that if the employer charges the premium $\pi^*(u_i, \hat{u}_{-i})$ then his expected utility is the same in both of the reported states of the world $(u_i, \hat{u}_{-i})$ and $(\hat{u}_i, \hat{u}_{-i})$.

If $\overline{\pi}(u_i) > \overline{\pi}(\hat{u}_i) > \overline{\pi}(\hat{u}_j)$ with j as depicted in Case 1 of Figure 1, then the optimal solution of (12) does not change, that is, $\pi^*(\hat{u}_i, \hat{u}_{-i}) = \pi^*(u_i, \hat{u}_{-i})$. Indeed, the only case for which the value of the objective function in (12) changes is when the employer sets the premium equal to $\overline{\pi}_m(\hat{u}_i, \hat{u}_{-i}) = \overline{\pi}(\hat{u}_i)$. Since $\overline{\pi}(\hat{u}_i) < \overline{\pi}(u_i)$, this change is negative.

If $\overline{\pi}(\hat{u}_j) > \overline{\pi}(\hat{u}_i) > \overline{\pi}(\hat{u}_k)$ with j and k as depicted in Case 1 of Figure 1, then the optimal solution can change. In this case the value of the objective function changes when the employer charges the premium $\overline{\pi}_m(\hat{u}_i, \hat{u}_{-i}) = \overline{\pi}(\hat{u}_i)$ or $\overline{\pi}_m(\hat{u}_i, \hat{u}_{-i}) = \overline{\pi}(\hat{u}_j)$. The first, however, cannot be optimal. To see this, consider the following two states, $(u_i, \hat{u}_{-i})$ and $(\hat{u}_i, \hat{u}_{-i})$. If the employer charges the premium $\overline{\pi}(\hat{u}_j)$ in state $(u_i, \hat{u}_{-i})$ and $\overline{\pi}(\hat{u}_i)$ in state $(\hat{u}_i, \hat{u}_{-i})$, then the same set of employees get the insurance. Since $\overline{\pi}(u_i) < \overline{\pi}(\hat{u}_j)$ this means that the employer's expected utility is higher in state $(u_i, \hat{u}_{-i})$ than in state $(\hat{u}_i, \hat{u}_{-i})$. So, if $\overline{\pi}(\hat{u}_j)$ would be the optimal premium in state $(\hat{u}_i, \hat{u}_{-i})$ then charging the premium $\overline{\pi}(\hat{u}_i)$ yields a higher expected utility than charging $\pi^*(u_i, \hat{u}_{-i})$. But then charging $\overline{\pi}(u_i)$ in state $(u_i, \hat{u}_{-i})$ must also yield a higher expected utility than $\pi^*(u_i, \hat{u}_{-i})$. This contradicts the optimality of $\pi^*(u_i, \hat{u}_{-i})$ in state $(u_i, \hat{u}_{-i})$. Hence, $\overline{\pi}(\hat{u}_i)$ cannot be optimal. A similar argument holds if the employer charges the premium $\overline{\pi}(\hat{u}_j)$. So, for the optimal premium $\pi^*(\hat{u}_i, \hat{u}_{-i})$ we have that $\pi^*(\hat{u}_i, \hat{u}_{-i}) = \pi^*(u_i, \hat{u}_{-i})$. Moreover, since $\overline{\pi}(\hat{u}_i) < \pi^*(\hat{u}_i, \hat{u}_{-i})$ it follows that employee i does not participate in the pension fund.

If $\overline{\pi}(hu_i) < \overline{\pi}(\hat{u}_k)$ than it follows from the same argument as above that $\overline{\pi}(\hat{u}_i) < \pi^*(\hat{u}_i, \hat{u}_{-i}) = \pi^*(u_i, \hat{u}_{-i})$.

Summarizing Case 1, if $\overline{\pi}(u_i) < \pi^*(u_i, \hat{u}_{-i})$ then $\overline{\pi}(\hat{u}_i) < \pi^*(\hat{u}_i, \hat{u}_{-i})$ if $\overline{\pi}(\hat{u}_i) < \overline{\pi}(u_i)$. So, if employee i is not allowed to partcipate in the pension fund when he reveals his true utility function u_i, then he also is not allowed to participate when he reports a utility function $\hat{u}_i$ with a lower maximum premium $\overline{\pi}(\hat{u}_i) < \overline{\pi}(u_i)$.

Case 2: In the second situation, that is, $\overline{\pi}(u_i) > \pi^*(u_i, \hat{u}_{-i})$, we have that employee i gets the insurance from his employer. Again, suppose that employee i reports a utility function $\hat{u}_i$ with $\overline{\pi}(\hat{u}_i) < \overline{\pi}(u_i)$.

If $\overline{\pi}(u_i) > \overline{\pi}(\hat{u}_i) \geq \overline{\pi}(\hat{u}_j)$ with j as depicted in Case 2 of Figure 1, then the optimum does not change. So, $\pi^*(\hat{u}_i, \hat{u}_{-i}) = \pi^*(u_i, \hat{u}_{-i})$.

If $\overline{\pi}(\hat{u}_j) > \overline{\pi}(\hat{u}_i) \geq \overline{\pi}(\hat{u}_k)$ with j and k as depicted in Case 2 of Figure 1, then the value of the objective function in (12) only changes if the employer charges the premium $\overline{\pi}_m(\hat{u}_i, \hat{u}_{-i}) = \overline{\pi}(\hat{u}_i)$ or $\overline{\pi}_m(\hat{u}_i, \hat{u}_{-i}) = \overline{\pi}(\hat{u}_j)$. Now, $\overline{\pi}(\hat{u}_i)$ cannot be an optimal premium. For if $\overline{\pi}(\hat{u}_i)$ is optimal in state $(\hat{u}_i, \hat{u}_{-i})$ then $\overline{\pi}(\hat{u}_j)$ must be optimal in state $(u_i, \hat{u}_{-i})$, which contradicts the optimality of $\pi^*(u_i, \hat{u}_{-i})$. A similar argument holds if the premium equals $\overline{\pi}(\hat{u}_j)$. Thus, for the optimal premium $\pi^*(\hat{u}_i, \hat{u}_{-i})$ we have that $\pi^*(\hat{u}_i, \hat{u}_{-i}) = \pi^*(u_i, \hat{u}_{-i})$.

If $\pi^*(u_i, \hat{u}_{-i}) \leq \overline{\pi}(\hat{u}_i) < \overline{\pi}(\hat{u}_k)$ then by the same argument as above it follows that $\pi^*(\hat{u}_i, \hat{u}_{-i}) = \pi^*(u_i, \hat{u}_{-i})$.

Summarizing Case 2, if $\pi^*(u_i, \hat{u}_{-i}) < \overline{\pi}(u_i)$ then $\pi^*(\hat{u}_i, \hat{u}_{-i}) = \pi^*(u_i, \hat{u}_{-i})$ for all $\hat{u}_i$ satisfying $\pi^*(u_i, \hat{u}_{-i}) \leq \overline{\pi}(\hat{u}_i) < \overline{\pi}(u_i)$. This means that employee i still participates in the pension fund if he reports a maximum premium of at least $\pi^*(u_i, \hat{u}_{-i})$ and at most $\overline{\pi}(u_i)$.

Case 3: In the third and final case we have that $\overline{\pi}(u_i) = \pi^*(u_i, \hat{u}_{-i})$. Suppose employee i steadily decreases his reported maximum premium and let $\overline{\pi}(\tilde{u}_i) < \overline{\pi}(u_i)$ be the first utility function for which $\overline{\pi}(\tilde{u}_i) \neq \pi^*(\tilde{u}_i, \hat{u}_{_i})$. Then either $\overline{\pi}(\tilde{u}_i) < \pi^*(\tilde{u}_i, \hat{u}_{-i})$ or $\overline{\pi}(\tilde{u}_i) > \pi^*(\tilde{u}_i, \hat{u}_{-i})$. If the first possibility occurs, we know from Case 1 that $\overline{\pi}(\hat{u}_i) < \pi^*(\hat{u}_i, \hat{u}_{-i})$ for all $\hat{u}_i$

with $\overline{\pi}(\hat{u}_i) < \overline{\pi}(\tilde{u}_i)$. If the latter possibility occurs then Case 2 applies. This means that $\pi^*(\hat{u}_i, \hat{u}_{-i}) = \pi^*(\tilde{u}_i, \hat{u}_i)$ for all $\hat{u}_i$ satisfying $\pi^*(\tilde{u}_i, \hat{u}_{-i}) \leq \overline{\pi}(\hat{u}_i) < \overline{\pi}(\tilde{u}_i)$.

Summarizing Case 3, if the premium π^* increases due to the fact employee i reports another, lower maximum premium $\overline{\pi}(\hat{u}_i)$, then this implies that employee i does not get any insurance when reporting $\overline{\pi}(\hat{u}_i)$, that is, $\overline{\pi}(\hat{u}_i) < \pi^*(\hat{u}_i, \hat{u}_{-i})$.

Now let us start with the second part of the proof. If the expected utility of employee i decreases when he reports $\hat{u}_i$ instead of u_i then one of the following two possibility applies: either he gets the insurance for a higher premium, or he gets the insurance when reporting u_i whereas he does not get it when reporting $\hat{u}_i$.

Let us begin with the first possibility, that is, he can still participate in the pension fund but for a higher premium. This means that $\overline{\pi}(u_i) \geq \pi^*(u_i, \hat{u}_{-i})$, $\overline{\pi}(\hat{u}_i) \geq \pi^*(\hat{u}_i, \hat{u}_{-i})$, and $\pi^*(u_i, \hat{u}_{-i}) < \pi^*(\hat{u}_i, \hat{u}_{-i})$. Since employee i gets the insurance when reporting u_i the cases 2 and 3 of Figure 1 apply. From the first part of this proof follows that, if $\overline{\pi}(u_i) \geq \pi^*(u_i, \hat{u}_{-i})$ and the optimal premium increases when $\overline{\pi}(\hat{u}_i)$ is reported instead of $\overline{\pi}(u_i)$ then $\overline{\pi}(\hat{u}_i) < \pi^*(\hat{u}_i, \hat{u}_{-i})$. This implies that employee i cannot partcipate in the pension fund. Hence, the case that employee i participates for a higher premium when deviating cannot occur.

The only possibility left for which the expected utility of employee i decreases is, if he gets the insurance when reporting u_i whereas he does not get it when reporting $\hat{u}_i$. This means that $\pi^*(u_i, \hat{u}_{-i}) \leq \overline{\pi}(u_i)$ and $\pi^*(\hat{u}_i, \hat{u}_{-i}) > \overline{\pi}(\hat{u}_i)$. Recall that when employee i is not allowed to participate in the pension fund if he reports $\hat{u}_i$, then his payoff $h_i^*(\hat{u}_i, \hat{u}_{-i})$ equals

$$\sum_{t=1}^{\min(P_i,N_i)} \delta^t (1-\tau) w_{it} \ + \sum_{t=P_i+1}^{N_i} \delta^t (1-\tau) y_{it},$$

or

$$\sum_{t=1}^{\min(P_i,N_i)} \delta^t ((1-\tau) w_{it} - \pi_i^m w_{it}) \ + \sum_{t=P_i+1}^{N_i} \delta^t (1-\tau)(y_{it} + E_{it}),$$

depending on which of the two outside options employee i prefers most, that is, no additional pension or one with an existing insurance company. From the definition of $\overline{\pi}(u_i)$ it then follows that $E(u_i(h_i^*(\hat{u}_i, \hat{u}_{-i})))$ equals the expected utility of employee i when he participates in the pension fund for his maximum premium $\overline{\pi}(u_i)$, that is

$$E(u_i(h_i^*(\hat{u}_i, \hat{u}_{-i}))) =$$

$$E\left(u_i\left(\sum_{t=1}^{\min(P_i,N_i)}\delta^t(1-\tau)(w_{it}-\overline{\pi}(u_i)w_{it}) \;+\; \sum_{t=P_i+1}^{N_i}\delta^t(1-\tau)(y_{it}+E_{it})\right)\right). \quad (13)$$

For ease of notation let us abbreviate the right hand side to $\nu(\overline{\pi}(u_i))$. Since $E(u_i(h_i^*(\hat{u}_i,\hat{u}_{-i}))) = \nu(\overline{\pi}(u_i))$ this also implies that employee i is indifferent between participating in the pension fund and not participating if $\pi^*(u_i,\hat{u}_{-i}) = \overline{\pi}(u_i)$. So, in that case, reporting $\hat{u}_i$ instead of u_i does not decrease his expected utility. Next, define

$$L(\hat{u}_i) = \{\hat{u}_{-i} \in \hat{U}_{-i} | \pi^*(u_i,\hat{u}_{-i}) < \overline{\pi}(u_i) \text{ and } \pi^*(\hat{u}_i,\hat{u}_{-i}) > \overline{\pi}(\hat{u}_i)\},$$

as the set of all $\hat{u}_{-i}$ for which revealing $\hat{u}_i$ instead of u_i strictly decreases the expected utility. Recall from the first part of the proof that $\pi^*(u_i,\hat{u}_{-i}) > \overline{\pi}(\hat{u}_i)$ for all $\hat{u}_{-i} \in L(\hat{u}_i)$.

For the third of the proof let us focus on the case that the expected utility increases. This is the case when either employee i participates in the pension fund for a lower premium, or employee i participates when reporting $\hat{u}_i$ whereas he cannot participate when reporting u_i.

Let us start with the latter possibility. This implies that $\pi^*(\hat{u}_i,\hat{u}_{-i}) \leq \overline{\pi}(\hat{u}_i) < \overline{\pi}(u_i) < \pi^*(u_i,\hat{u}_{-i})$. Since $\overline{\pi}(u_i) < \pi^*(u_i,\hat{u}_{-i})$ Case 1 applies. From the first part of the proof we know that if $\overline{\pi}(u_i) < \pi^*(u_i,\hat{u}_{-i})$ and $\overline{\pi}(\hat{u}_i) < \overline{\pi}(u_i)$ then also $\overline{\pi}(\hat{u}_i) < \pi^*(\hat{u}_i,\hat{u}_{-i})$. So, employee i can also not participate in the pension fund when he reports $\hat{u}_i$. Thus, the case that employee i does not participate in the pension fund when reporting u_i but does participate when reporting $\hat{u}_i$ cannot occur.

The remaining possibility is the one where employee i gets the insurance for a lower premium. Thus, $\pi^*(\hat{u}_i,\hat{u}_{-i}) \leq \overline{\pi}(\hat{u}_i)$, $\pi^*(u_i,\hat{u}_{-i}) \leq \overline{\pi}(u_i)$, and $\pi^*(\hat{u}_i,\hat{u}_{-i}) < \pi^*(u_i,\hat{u}_{-i})$. In fact, to show that h^* is not incentive compatible it is sufficient to consider only those cases where $\pi^*(\hat{u}_i,\hat{u}_{-i}) \leq \overline{\pi}(\hat{u}_i) < \overline{\pi}(u_i) = \pi^*(u_i,\hat{u}_{-i})$. Let $W(\hat{u}_i)$ denote the set of all $\hat{u}_{-i}$ satisfying this condition, that is,

$$W(\hat{u}_i) = \{\hat{u}_{-i} \in \hat{U}_{-i} | \pi^*(\hat{u}_i,\hat{u}_{-i}) \leq \overline{\pi}(\hat{u}_i) < \overline{\pi}(u_i) = \pi^*(u_i,\hat{u}_{-i})\}.$$

Then the expected utility of employee i increases when reporting $\hat{u}_i$ instead of u_i if $\hat{u}_{-i} \in W(\hat{u}_i)$. Moreover, since $\pi^*(u_i,\hat{u}_{-i}) = \overline{\pi}(u_i)$ it holds that $E(u_i(h_i^*(u_i,\hat{u}_{-i}))) = \nu(\overline{\pi}(u_i))$ for all $\hat{u}_{-i} \in W(\hat{u}_i)$.

Now, let us turn to the fourth part of the proof. If the payoff function h^* satisfies incentive compatibility, then for the true utility function u_i we must have that

$$\int_{U_{-i}} E(u_i(h_i^*(u_i, \hat{u}_{-i})))d\mu_i(u_i) \geq \int_{U_{-i}} E(u_i(h_i^*(\hat{u}_i, \hat{u}_{-i})))d\mu_i(u_i),$$

for all $\hat{u}_i \in U_i$. In particular, this holds for $\hat{u}_i \in U_i$ such that $\overline{\pi}(\hat{u}_i) < \overline{\pi}(u_i)$. Note that deviating to $\hat{u}_i$ decreases the expected utility of employee i when $\hat{u}_{-i} \in L(\hat{u}_i)$ and increases the expected utility when $\hat{u}_{-i} \in W(\hat{u}_i)$. So, if deviating to $\hat{u}_i$ is not profitable, then the total loss incurred for all $\hat{u}_{-i} \in L(\hat{u}_i)$ must certainly exceed the total profits obtained for all $\hat{u}_{-i} \in W(\hat{u}_i)$. This implies that

$$\begin{aligned}\int_{L(\hat{u}_i)} & E(u_i(h_i^*(u_i, \hat{u}_{-i}))) \; - \; E(u_i(h_i^*(\hat{u}_i, \hat{u}_{-i})))d\mu_i(u_i) \geq \\ & \int_{W(\hat{u}_i)} E(u_i(h_i^*(\hat{u}_i, \hat{u}_{-i}))) \; - \; E(u_i(h_i^*(u_i, \hat{u}_{-i})))d\mu_i(u_i)\end{aligned}$$

for all $\hat{u}_i \in U_i$ with $\overline{\pi}(\hat{u}_i) < \overline{\pi}(u_i)$. Since $\hat{u}_{-i} \in L(\hat{u}_i)$ we know from expression (13) that $E(u_i(h_i^*(\hat{u}_i, \hat{u}_{-i}))) = \nu(\overline{\pi}(u_i))$. Furthermore, since $\hat{u}_{-i} \in W(\hat{u}_i)$ implies that $\pi^*(u_i, \hat{u}_{-i}) = \overline{\pi}(u_i)$, it also holds that $E(u_i(h_i^*(u_i, \hat{u}_{-i}))) = \nu(\overline{\pi}(u_i))$. Hence, if h^* is incentive compatible then

$$\begin{aligned}\int_{L(\hat{u}_i)} & E(u_i(h_i^*(u_i, \hat{u}_{-i}))) \; - \; \nu(\overline{\pi}(u_i))d\mu_i(u_i) \geq \\ & \int_{W(\hat{u}_i)} E(u_i(h_i^*(\hat{u}_i, \hat{u}_{-i}))) \; - \; \nu(\overline{\pi}(u_i))d\mu_i(u_i)\end{aligned} \tag{14}$$

for all $\hat{u}_i \in U_i$ with $\overline{\pi}(\hat{u}_i) < \overline{\pi}(u_i)$. Since $\hat{u}_{-i} \in L(\hat{u}_i)$ implies that $\pi^*(u_i, \hat{u}_{-i}) > \overline{\pi}(\hat{u}_i)$ we have that

$$E(u_i(h_i^*(u_i, \hat{u}_{-i}))) < \nu(\overline{\pi}(\hat{u}_i)). \tag{15}$$

Similarly, since $\hat{u}_{-i} \in W(\hat{u}_i)$ implies that $\pi^*(\hat{u}_i, \hat{u}_{-i}) \leq \overline{\pi}(\hat{u}_i)$ it follows that $E(u_i(h_i^*(\hat{u}_i, \hat{u}_{-i}))) \geq \nu(\overline{\pi}(\hat{u}_i))$. Thus if inequality (14) holds for all $\hat{u}_i \in U_i$ with $\overline{\pi}(\hat{u}_i) < \overline{\pi}(u_i)$, it must also hold true that

$$\begin{aligned}\int_{L(\hat{u}_i)} & \nu(\overline{\pi}(\hat{u}_i)) \; - \; \nu(\overline{\pi}(u_i))d\mu_i(u_i) \geq \\ & \int_{W(\hat{u}_i)} \nu(\overline{\pi}(\hat{u}_i)) \; - \; \nu(\overline{\pi}(u_i))d\mu_i(u_i).\end{aligned}$$

Since $\overline{\pi}(\hat{u}_i) < \overline{\pi}(u_i)$ implies that $\nu(\overline{\pi}(\hat{u}_i)) - \nu(\overline{\pi}(u_i)) > 0$ we get that

$$\int_{L(\hat{u}_i)} d\mu_i(u_i) \; \geq \; \int_{W(\hat{u}_i)} d\mu_i(u_i),$$

or, equivalently, that $\mu_i(u_i)(L(\hat{u}_i)) \geq \mu_i(u_i)(W(\hat{u}_i))$ for all $\hat{u}_i \in U_i$ with $\overline{\pi}(\hat{u}_i) < \overline{\pi}(u_i)$.

Summarizing, if h^* is incentive compatible then for all $i \in N$ and all $u_i, \hat{u}_i \in U_i$ such that $\overline{\pi}(\hat{u}_i) < \overline{\pi}(u_i)$ it holds that $\mu_i(u_i)(L(\hat{u}_i)) \geq \mu_i(u_i)(W(\hat{u}_i))$.

For the fifth and final step of the proof we need the following two lemmas. The first lemma shows that $L(\hat{u}_i)$ is decreasing in $\overline{\pi}(\hat{u}_i)$. The second lemma shows that $W(\hat{u}_i)$ is increasing in $\overline{\pi}(\hat{u}_i)$.

Lemma 5.1 Let $\hat{u}_i, \tilde{u}_i, u_i \in U_i$ be such that $\overline{\pi}(\hat{u}_i) < \overline{\pi}(\tilde{u}_i) < \overline{\pi}(u_i)$. Then $L(\tilde{u}_i) \subset L(\hat{u}_i)$.

PROOF: Take $\hat{u}_{-i} \in L(\tilde{u}_i)$. Then $\overline{\pi}(\tilde{u}_i) < \pi^*(\tilde{u}_i, \hat{u}_{-i})$ and $\overline{\pi}(u_i) > \pi^*(u_i, \hat{u}_{-i})$. From $\overline{\pi}(\tilde{u}_i) < \pi^*(\tilde{u}_i, \hat{u}_i)$ it follows that Case 1 - discussed in the first part of the proof - applies. Since $\overline{\pi}(\hat{u}_i) < \overline{\pi}(\tilde{u}_i)$ this implies that $\overline{\pi}(\hat{u}_i) < \pi^*(\hat{u}_i, \hat{u}_{-i})$. Hence, $\hat{u}_{-i} \in L(\hat{u}_i)$. $\square$

Lemma 5.2 Let $\hat{u}_i, \tilde{u}_i, u_i \in U_i$ be such that $\overline{\pi}(\hat{u}_i) < \overline{\pi}(\tilde{u}_i) < \overline{\pi}(u_i)$. Then $W(\hat{u}_i) \subset W(\tilde{u}_i)$.

PROOF: Take $\hat{u}_i \in W(\hat{u}_i)$. So, $\overline{\pi}(u_i) = \pi^*(u_i, \hat{u}_{-i}) > \overline{\pi}(\hat{u}_i) \geq \pi^*(\hat{u}_i, \hat{u}_{-i})$. Since $\overline{\pi}(u_i) = \pi^*(u_i, \hat{u}_{-i})$ it follows that Case 3 - discussed in the first part of the proof - applies. Hence, from $\overline{\pi}(\tilde{u}_i) > \overline{\pi}(\hat{u}_i) \geq \pi^*(\hat{u}_i, \hat{u}_{-i})$ it follows that $\overline{\pi}(\tilde{u}_i) \geq \pi^*(\tilde{u}_i, \hat{u}_{-i})$. Then $\pi^*(u_i, \hat{u}_{-i}) = \overline{\pi}(u_i) > \overline{\pi}(\tilde{u}_i)$ implies that $\pi^*(u_i, \hat{u}_{-i}) > \pi^*(\tilde{u}_i, \hat{u}_{-i})$. Thus, $\hat{u}_{-i} \in W(\tilde{u}_i)$. $\square$

Next, consider a sequence $(\hat{u}_i^k)_{k \in \mathbb{N}}$ such that $\overline{\pi}(\hat{u}_i^k) < \overline{\pi}(\hat{u}_i^{k+1})$ for all $k \in \mathbb{N}$ and $\lim_{k \to \infty} \overline{\pi}(\hat{u}_i^k) = \overline{\pi}(u_i)$. Then for all $k \in \mathbb{N}$ it holds that $\mu_i(u_i)(L(\hat{u}_i^k)) \geq \mu_i(u_i)(W(\hat{u}_i^k))$. Since $0 \leq \mu_i(u_i)(L(\hat{u}_i^k)) \leq 1$ for all $k \in \mathbb{N}$ and $L(\hat{u}_i^1) \supset L(\hat{u}_i^2) \supset \ldots$ it follows that $\left(\mu_i(u_i)(L(\hat{u}_i^k))\right)_{k \in \mathbb{N}}$ is a decreasing convergent sequence with limit equal to $\mu_i(u_i)(L)$, where $L = \cap_{k \in \mathbb{N}} L(\hat{u}_i^k)$. Similarly, since $0 \leq \mu_i(u_i)(W(\hat{u}_i^k)) \leq 1$ for all $k \in \mathbb{N}$ and $W(\hat{u}_i^1) \subset W(\hat{u}_i^2) \subset \ldots$ it follows that $\left(\mu_i(u_i)(W(\hat{u}_i^k))\right)_{k \in \mathbb{N}}$ is an increasing convergent sequence with limit equal to $\mu_i(u_i)(W)$, where $W = \cup_{k \in \mathbb{N}} W(\hat{u}_i^k)$. Moreover, $\mu_i(u_i)(L) \geq \mu_i(u_i)(W)$.

Now it holds that $L = \cap_{k \in \mathbb{N}} L(\hat{u}_i^k) = \emptyset$. To see this, recall that

$$L(\hat{u}_i^k) = \{\hat{u}_{-i} \in \hat{U}_{-i} | \pi^*(u_i, \hat{u}_{-i}) < \overline{\pi}(u_i) \text{ and } \pi^*(\hat{u}_i^k, \hat{u}_{-i}) > \overline{\pi}(\hat{u}_i^k)\},$$

and that $\pi^*(u_i, \hat{u}_{-i}) > \overline{\pi}(\hat{u}_i^k)$ for all $\hat{u}_{-i} \in L(\hat{u}_i^k)$. So, if $\hat{u}_{-i} \in L$ then $\pi^*(u_i, \hat{u}_{-i}) > \overline{\pi}(\hat{u}_i^k)$ for all $k \in \mathbb{N}$. Since $\lim_{k \to \infty} \overline{\pi}(\hat{u}_i^k) = \overline{\pi}(u_i)$ this implies that $\pi^*(u_i, \hat{u}_{-i}) \geq \overline{\pi}(u_i)$. Obviously, this contradicts $\pi^*(u_i, \hat{u}_{-i}) < \overline{\pi}(u_i)$. Hence, $L = \emptyset$ so that $0 = \mu_i(u_i)(L) \geq \mu_i(u_i)(W) \geq 0$. Consequently, h^* is not incentive compatible if we can show that $\mu_i(u_i)(W) > 0$.

In order to show that $\mu_i(u_i)(W) > 0$ for some $i \in N$, take $\hat{u} \in U$ such that $\overline{\pi}(\hat{u}_j) \neq \overline{\pi}(\hat{u}_k)$ for all $j, k \in N$ with $j \neq k$ and such that the optimum premium $\pi^*(\hat{u})$ is unique. Take $i \in N$ such that $\overline{\pi}(\hat{u}_i) = \pi^*(\hat{u})$. So, $\overline{\pi}(\hat{u}_i)$ is the premium charged by the employer. Since $\overline{\pi}(\hat{u}_j) \neq \overline{\pi}(\hat{u}_k)$ for all $j, k \in N$ with $j \neq k$, there exists a utility function $u_i \in U_i$ satisfying

(i) $\pi^*(u_i, \hat{u}_{-i}) = \overline{\pi}(u_i)$, that is, $\overline{\pi}(u_i)$ is the premium that the employer charges for the insurance when employee i reveals u_i instead of $\hat{u}_i$;

(ii) $\overline{\pi}(u_i) > \overline{\pi}(\hat{u}_i)$;

(iii) for all $j \in N$ with $\overline{\pi}(\hat{u}_j) > \overline{\pi}(\hat{u}_i)$ it holds that $\overline{\pi}(\hat{u}_j) > \overline{\pi}(u_i)$.

Thus the maximum premium $\overline{\pi}(u_i)$ is chosen such that it is slightly greater than $\overline{\pi}(\hat{u}_i)$ and such that $\overline{\pi}(u_i)$ is still optimal in state $(u_i, \hat{u}_{-i})$. So, we have that $\pi^*(\hat{u}_i, \hat{u}_{-i}) = \overline{\pi}(\hat{u}_i)$ and $\pi^*(u_i, \hat{u}_{-i}) = \overline{\pi}(u_i)$.

Now, suppose that u_i is employee i's true utility function and that the other employees report $\hat{u}_{-i}$. Since the premium decreases from $\overline{\pi}(u_i)$ to $\overline{\pi}(\hat{u}_i)$ if employee i reveals $\hat{u}_i$ instead of u_i, employee i benefits from deviating. Hence, $\hat{u}_{-i} \in W(\hat{u}_i)$.

Next, consider employee $j \neq i$ and let employee i report u_i. If $\overline{\pi}(\hat{u}_j) > \overline{\pi}(u_i)$ then $\overline{\pi}(u_i)$ is still optimal if employee j reports any utility function v_j satisfying $\overline{\pi}(u_i) < \overline{\pi}(v_j) \leq \overline{\pi}(\hat{u}_j)$. If $\overline{\pi}(u_i) > \overline{\pi}(\hat{u}_j)$ then $\overline{\pi}(u_i)$ is still optimal if employee j reports any utility function v_j satisfying $\overline{\pi}(v_j) < \overline{\pi}(\hat{u}_j)$. The same result holds if employee i reports $\hat{u}_i$.

Since employee j is risk averse we have that the maximum premium $\overline{\pi}(v_j)$ has some lower bound l_j, for example, the maximum premium he would be willing to pay as a risk neutral employee. Define

$$B = \left\{ v_{-i} \in \hat{U}_{-i} \;\middle|\; \begin{array}{ll} \overline{\pi}(\hat{u}_i) < \overline{\pi}(v_j) \leq \overline{\pi}(\hat{u}_j), & \text{if } \overline{\pi}(\hat{u}_j) > \overline{\pi}(u_i) \\ l_j < \overline{\pi}(v_j) \leq \overline{\pi}(\hat{u}_j), & \text{if } \overline{\pi}(\hat{u}_j) < \overline{\pi}(u_i) \end{array} \right\}.$$

Since $\pi^*(u_i, v_{-i}) = \overline{\pi}(u_i)$ and $\pi^*(\hat{u}_i, v_{-i}) = \overline{\pi}(\hat{u}_i)$, employee i benefits from reporting $\hat{u}_i$ instead of u_i when $v_{-i} \in B$. Hence, $B \subset W(\hat{u}_i)$. Moreover, $\{\overline{\pi}(v_{-i}) | v_{-i} \in B\}$ is a product of intervals with strictly positive Lebesgue measure. Thus, $\mu_i(u_i)(B) > 0$ so that $\mu_i(u_i)(W(\hat{u}_i)) > 0$.

Now, let $(\hat{u}_i^k)_{k \in \mathrm{N}}$ be a sequence in U_i such that $\hat{u}_i^1 = \hat{u}_i$, $\overline{\pi}(\hat{u}_i^{k+1}) > \overline{\pi}(\hat{u}_i^k)$ for all $k \in \mathbb{N}$, and $\lim_{k \to \infty} \overline{\pi}(\hat{u}_i^k) = \overline{\pi}(u_i)$. Then $\mu_i(u_i)(W) \geq \mu_i(u_i)(W(\hat{u}_i^1)) > 0$, which implies that h^* is not incentive compatible.

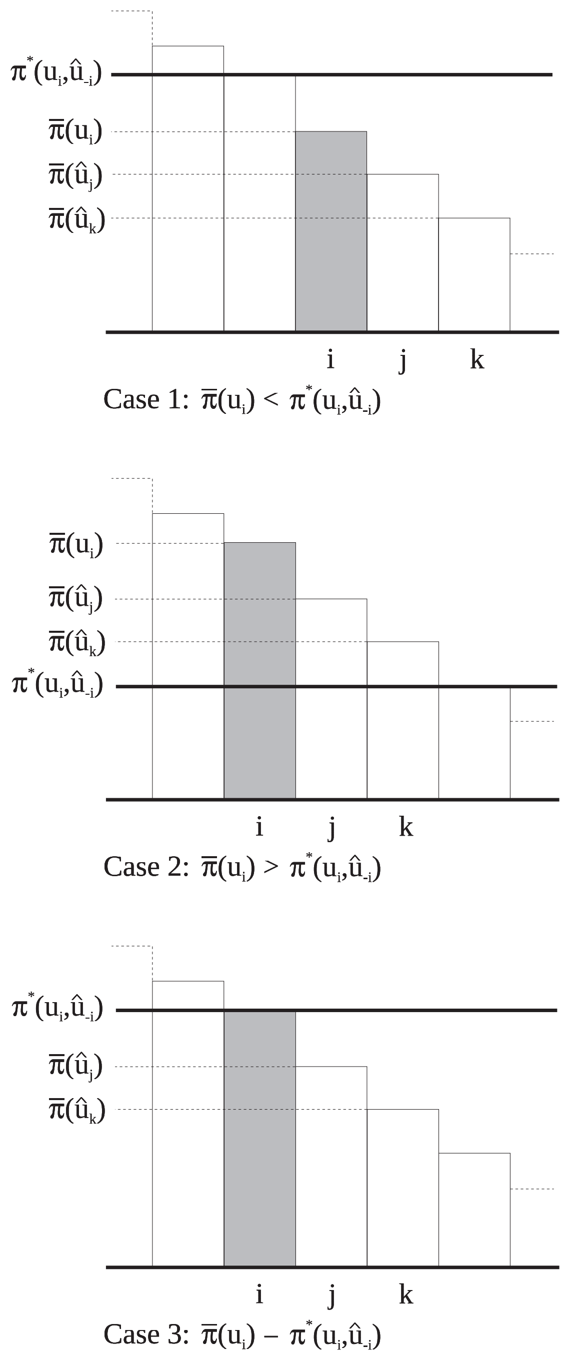

FIGURE 1: The relation between the reported premium and the optimal premium.

References

JACKSON, M.: "Bayesian Implementation," *Econometrica*, **59** (1991), 461-477.

MOORE, J.: "Implementation, Contracts, and Renegotiation in Environments with Complete Information," in *Advances in Economic Theory*, Vol. I, ed. J. Laffont. Cambridge: Cambridge University Press, 1992.

PALFREY, T.: "Implementation in Bayesian Equilibrium: the Multiple Equilibrium Problem in Mechanism Design," in *Advances in Economic Theory*, Vol. I, ed. J. Laffont. Cambridge: Cambridge University Press, 1992.